

Patent Abstracts of Japan

PUBLICATION NUMBER : 2000172722
PUBLICATION DATE : 23-06-00

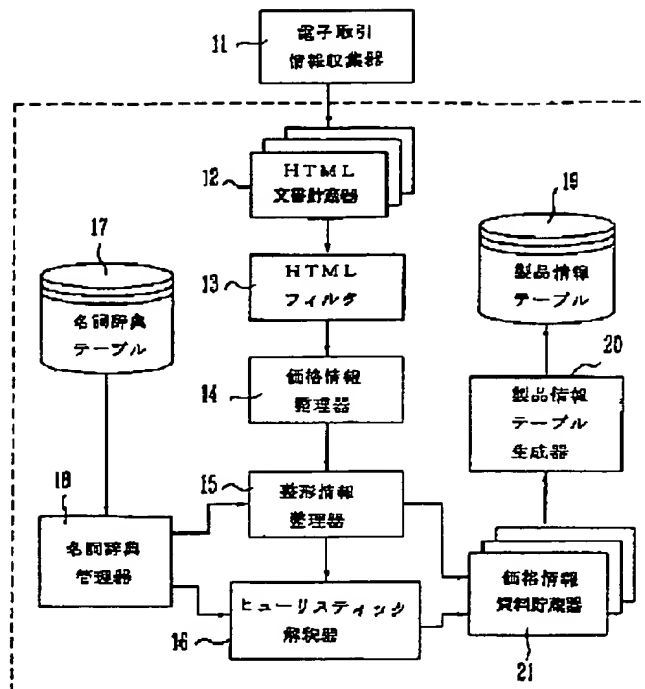
APPLICATION DATE : 29-11-99
APPLICATION NUMBER : 11338623

APPLICANT : KOREA ELECTRONICS
TELECOMMUN;

INVENTOR : BOKU SOHO;

INT.CL. : G06F 17/30 G06F 17/60

TITLE : METHOD AND SYSTEM FOR
PRODUCT INFORMATION AUTOMATIC
INDEXING OF ON-LINE STORE



ABSTRACT : PROBLEM TO BE SOLVED: To enable a comparison purchase at an on-line store on the web by automatically extracting a product information record through a heuristic interpreter so that product information tends to be positioned in a hypertext mark-up language(HTML) document.

SOLUTION: Unnecessary HTML documents among gathered HTML documents are filtered off primarily by an HTML filter 13 and a price information rearranging unit 14 performs conversion to a form suitable for use on a product information extraction subsystem. A formation information rearranging unit 15 decides the type of an input document and calls an analytic module for a type which is already analyzed to extract product information. A document having price information which failed to be analyzed is passed to the heuristic interpreter 16, which extracts product information. Although there is price information based upon product information and price information from a noun dictionary table 17, a document which does not have its product information extracted by the formation information rearranging device 15 is sequentially heuristically processed to extract the product information.

COPYRIGHT: (C)2000,JPO